International Livestock Research Institute

Training course report

Ecological niche modelling

September 2014

Written by Gladys Mosomtai

Edited by Bernard Bett and Tezira Lore

Citation
ILRI (International Livestock Research Institute). 2014. *Ecological niche modelling: Report of a training course held at Nairobi, Kenya on 15-19 September 2014.* Nairobi, Kenya: ILRI.

Contents

Project background

In poor countries, diseases associated with agriculture have important human health impacts. Food that nourishes can also sicken and kill. The Food Safety and Zoonoses program of the International Livestock Research Institute (ILRI) focuses on assessing and managing the health risks associated with food safety, zoonotic diseases (transmissible between people and animals) emerging diseases and diseases associated with agricultural intensification.

A major focus of our research is approaches to food safety and disease control that can work in informal markets and marginal areas. The program combines risk analysis, epidemiology, socio-economics and gender analysis to identify, test and adapt appropriate options for food safety assurance and control of zoonotic diseases within the context of developing countries.

Program activities are directly aligned with one of the four components of the CGIAR Research Program on Agriculture for Nutrition and Health, namely, prevention and control of agriculture-associated diseases. It also contributes to the value chain work of the CGIAR Research Program on Livestock and Fish.

Training summary

Organizer/co-organizers: Bernard Bett, Gladys Mosomtai and Rosekellen Njiru
Lecturers/facilitators: Gladys Mosomtai, Thomas Gachie, A. Townsend Peterson, Chris Yesson and Dave Redding

Ecological niche modelling has fast become an extremely popular tool in the use, exploration and analysis of biodiversity data. It offers the possibility of understanding species' geographic distributions (actual and potential) in detail, as well as species' environmental requirements (i.e. the ecological niche).

Although tools for niche modelling are now widely and freely available, a clear and appropriate thinking framework for the use of the tool is often not at hand, leading to many partial or inappropriate implementations. This course is designed to provide in-depth training in the use of niche modelling as an approach to biodiversity data.

The participants targeted are post doc fellows, PhD students and members of ILRI and government institutions who work in the areas of natural resource management, biodiversity, disease control and spatial mapping. The course was conducted at ILRI's Nairobi campus on 15-19 September 2014.

Agenda

Introduction of the concept of ecological niche modelling

Participants were introduced to the basic concepts of ecological niche modelling. The underpinning theories of realized niche verse occupied niche, fundamental niche as defined by environmental space and how it is projected to geographic space. Emphasis was made on how to collect species occurrence data which is one of the inputs, especially presence only data, since the results will depend wholly on the occurrence data. If the sampled occurrence data is not representative of all the possible niches the species occupies, then the model results might not be accurate. Possible errors that might occur during data collection were also discussed, for example, GPS errors, conversion errors from degree-minutes-seconds to decimal degree, taxonomic errors from misidentifying the species of interest and errors from labelling in cases like museum collections where no GPS co-ordinates were taken. Participants were shown methods in which they can correct this errors before running their models.

Environmental variables that define the niche of the species was expounded. Sources of this data was explained such as remotely sensed data, interpolated data, already available data that can be downloaded such as WorldClim data and the format that they should be when used in modelling. Participants raised issues regarding spatial resolution in which environmental variables should have and there was no thumb rule for this question, it wholly depends on the researcher and at what scale is he/she looking at the species of interest, at a local scale, it is advisable to use data with higher resolution compared to looking at a species at a regional scale. Emphasis on the concepts and data input for ecological niche modelling was later done by one of the guest speaker Prof. Townsend Peterson via google hangouts and due to time limit participants requested that he may give his lecture another day again.

Model algorithms were explained in detailed and the assumptions made in various algorithms and how this assumptions relate to the occurrence data collected in the field. Participants were made to understand why different algorithms were preferred over others especially when dealing with data like species occurrence which is binary and how the various algorithms deals with environmental variables for instance when there is high correlation or when number of environmental variables is small, also when sample size is small. The following algorithms were expounded:

- regression methods such as generalized linear models, generalized additive models, multivariate adaptive regression splines
- machine learning methods such as classification and regression trees, boosted regression trees random forest and Maximum Entropy (MaxEnt)
- profile methods such as Bioclim, Domain, Mahalanobis

Dr Chris Yesson, who was a guest speaker, also gave a lecture that expounded on the various algorithms and gave examples of work done using the different algorithms.

Practical exercises

Participants had an opportunity to do a hands-on practical exercise using MaxEnt and OpenModeller software. MaxEnt has a very friendly and easy-to-use interface compared to GARP which is one of the many algorithms assembled in OpenModeller. Participants preferred to use MaxEnt. Outputs were discussed by all the participants and how to interpret them. MaxEnt results that come in form of an HTML file were found to be really informative. Participants were also introduced to the basics of geographic information systems and a practical exercise was done using QGIS which is an open-source software. Participants were able to view their results in QGIS and visualize the niche of Rift Valley fever outbreaks based on centroids occurrence data. Furthermore, participants were introduced to the basics of R software and ran the DISMO package that is used to do species distribution in R.

Learning how to do model evaluation

Participants learned how to do model evaluation and a detailed explanation on the confusion matrix was given, based on the possible outcomes that may result from the model prediction, namely, true presence, true absence, false presence and false absence. A good model is one that is able to distinguish absence and presence correctly. Area under the curve (AUC) is mostly used in ecological niche modelling as the measure of accuracy. A detailed explanation was made on how AUC is generated which is a threshold independent measure of accuracy. Threshold independent accuracy was also explained such as Kappa statistics, correct classification rate and true skills statistics. Other model evaluation methods were also incorporated by the participants such as AIC and BIC and the guest speakers emphasized that AUC of the receiver operating characteristic (ROC) curve should not be the only measure of accuracy.

We introduced participants to another method of evaluating their models. This was after scientists found that AUC had weaknesses when it comes to ecological niche modelling. For instance, various algorithms generate the AUC curve differently. For instance, MaxEnt algorithm generates the curve from the origin while an algorithm like GARP generates the curve from 60% of the x-axis, hence AUC of MaxEnt will always be higher than that of GARP despite the fact that they will predict the same area. To solve this problem, Prof. Townsend (who was one of the guest speakers) and his team developed partial ROC which evaluates a part of the curve where all the algorithms measure and then generated a software known as Partial ROC (pROC). The participants were trained on how to use pROC so that they can make informed decisions based on their model outputs.

Recommendations

After the workshop, participants formed working groups to model Rift Valley fever, leptospirosis and anthrax across Kenya, Tanzania and Uganda. Three task groups were formed to steer the working groups and tentative deadlines were made, with Rift Valley fever modelling being the earliest due to the availability of data. Leptospirosis and anthrax would take a while due to unavailability of data. The groups agreed to meet again at another workshop in February (tentatively) to present their progress and invite other stakeholders like Kenya Wildlife Service who were not present at the workshop.

Training material

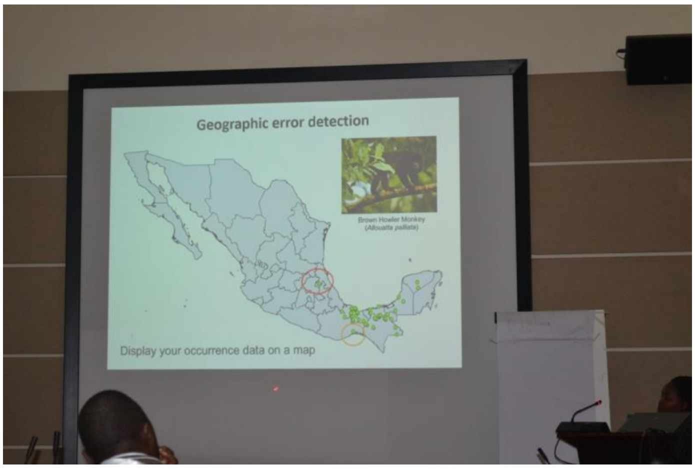

Photo 1: A PowerPoint presentation was used to train the participants. The presentation was shared with all the participants.

Photo 2: Practical exercise that involved the participants using the software.

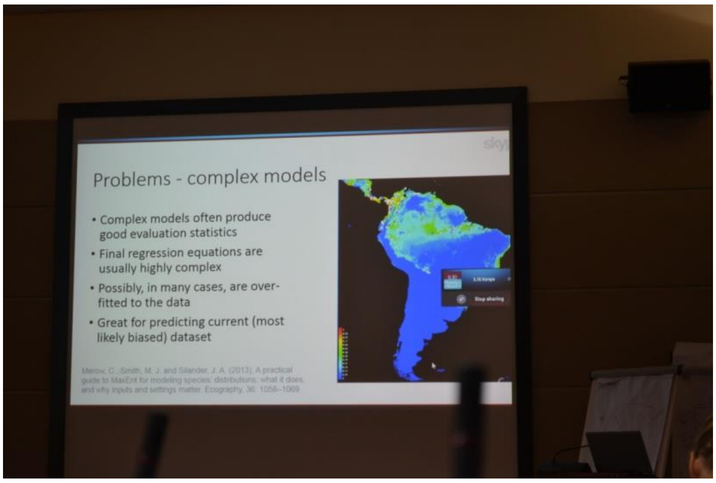

Photo 3: Skype presentations.

List of participants

Serial No.	Name	Email contact	Sex (M/F)	Country of origin	Country Classification (Developing/Developed)
1	Mark Nanyingi	mnanyingi@gmail.com	M	Kenya	Developing
2	Peninah Munyua	PMunyua@ke.cdc.gov	F	Kenya	Developing
3	Eric Osoro	osoroe@yahoo.com	M	Kenya	Developing
4	Jusper Kiplimo	j.kiplimo@cgiar.org	M	Kenya	Developing
5	Audi Allan		M	Kenya	Developing
6	Luke Nyakarahuka	nyakarahuka@gmail.com	M	Uganda	Developing
7	Austine Bitek	bitekorinde@yahoo.com	M	Kenya	Developing
8	Calvin Sindato	kndato@yahoo.co.uk	M	Tanzania	Developing
9	Salome Wanyoike	swwanyoike@yahoo.com	F	Kenya	Developing
10	Johanna Lindahl	j.lindahl@cgiar.org	F	Sweden	Developed
11	Susan Njenga	s.njenga@cgiar.org	F	Kenya	Developing
12	Purity Kiunga	Pkiunga@yahoo.com	F	Kenya	Developing
13	Joel Lutomiah		M	Kenya	Developing
14	Christine Atherstone	christine.atherstone@gmail.com	F	USA	Developed
15	Bernard Bett	b.bett@cgiar.org	M	Kenya	Developing
16	Sam Thumbi		M	Kenya	Developing